DOI: 10.54254/2754-1169/67/20241290

Trend Forecast of Shanghai Stock Exchange Composite Index Based on Monetary Supply and Consumer Price Index

Yi Peng[1,a,*], Yingji Xu[1,b]

[1]Minhang Crosspoint Academy at Shanghai Wenqi Middle School, Shanghai China
a. m13799346475@163.com, b. jayxu0215@163.com
**corresponding author*

Abstract: Stock market indices often serve as indicators of a country's economic conditions. Therefore, analysing the trends of stock market indices can assist individuals, institutions, and even governments in comprehending the state of the economy and developing suitable investment strategies or economic policies. However, accurately predicting these indices poses a significant challenge. In recent years, machine learning has displayed remarkable learning capabilities in various industries, making it an intriguing and viable avenue for trend prediction. In this article, we have selected two closely linked data sources, namely monetary supply and consumer price index, which are highly correlated with economic operations. By combining these data with regression models, we have developed an algorithm for predicting China's Shanghai Stock Exchange Composite Index (SSECI). Experimental results illustrate a strong correlation between the collected data and the index, highlighting their value in indicating economic conditions.

Keywords: Stock Index Prediction, Linear Regression, AdaBoost

1. Introduction

Stock market indices, such as Standard & Poor’s 500 and Nasdaq 100, serve as indicators of a country's economic conditions [1]. Understanding the trend of these indices is crucial for individuals, institutions, and even governments to gain insight into the state of the economy and formulate appropriate investment strategies or economic policies. Despite international skepticism due to China's government foreign exchange controls and financial policies, it is important to determine if China's stock market, particularly the Shanghai Stock Exchange Composite Index (SSECI), is truly unpredictable. This study aims to investigate any potential clues in China's stock market.

The stock market is a complex and dynamic system that is inherently unstable, nonlinear, and influenced by various factors, including the future economy and the international situation. Thus, it is imperative to identify relevant factors that reflect the economic state and select appropriate prediction algorithms. In this study, we have identified two data sources closely related to the economic situation: the Consumer Price Index (CPI) [2] and the Monetary Supply. We have chosen these factors for the following reasons:

- CPI: The CPI serves as a measure of inflation, and changes in inflation expectations can impact the stock market. If investors anticipate high inflation, they may demand higher returns on their investments, leading to a decrease in stock prices. Conversely, low inflation expectations can be positive for stocks as it supports stronger purchasing power for consumers.

DOI: 10.54254/2754-1169/67/20241290

Moreover, changes in CPI can influence central banks' decisions regarding interest rates. A rise in CPI may prompt central banks to increase interest rates to control inflation. Higher interest rates can increase borrowing costs for businesses and consumers, potentially impacting corporate profits and consumer spending, which in turn can affect stock market performance.

- Monetary Supply: An increase in monetary supply can result in higher liquidity in the economy. Ample liquidity makes it easier for investors to buy and sell stocks, potentially increasing trading volumes and market activity. This increased liquidity can have a positive impact on the stock market by providing investors with more trading opportunities.

Additionally, changes in monetary supply can influence interest rates set by central banks. An increase in monetary supply may lead central banks to lower interest rates in an effort to stimulate economic activity. Lower interest rates can make borrowing cheaper, potentially boosting corporate profits and increasing stock market performance. Conversely, a decrease in monetary supply may result in higher interest rates, negatively impacting the stock market as borrowing costs increase.

Furthermore, an increase in monetary supply can contribute to inflationary pressures in the economy. If investors anticipate higher inflation, they may demand higher returns on their investments, leading to a decrease in stock prices. Conversely, a scarcity of currency supply can lead to deflationary pressures, which may negatively impact the stock market.

It should be noted that the data for these factors cannot be collected on a daily or weekly basis, so our analysis focuses on a monthly frequency. The challenge lies in utilizing these two factors to predict the trend of the SSECI. A straightforward approach is to apply Linear Regression (LR) [3] to predict the SSECI value each month. However, due to the complexity of the problem, a simple LR model is insufficient. According to the PAC theory [4], which suggests that an ensemble of weak learners can create a strong learner, we turn to use AdaBoost [5] to address this task. Specifically, we combine the CPI, Currency Supply, and SSECI at month t to create a feature vector and use the SSECI of the next month, t+1, as the ground truth.

Experimental results demonstrate the efficiency of these two factors in predicting the trend of the SSECI. Furthermore, we develop an investment strategy based on the AdaBoost Regressor, which yields an average profit ratio of over 10% in simulation. In summary, our contributions are as follows:

- We are the first to investigate the relationship between the CPI and the Currency Supply with the Shanghai Stock Exchange Composite Index, and apply these factors in predicting the index's trend using AdaBoost.
- We collect data on the CPI and the Monetary Supply from China's government statistics from 2016 to 2023 and record the closing price of SSECI each month. Experiments on this customized dataset demonstrate a close correlation between these two factors and the SSECI.

2. Data Description

In this section, we provide details on the two data sources we have identified. We collected the Consumer Price Index (CPI) and Monetary Supply data from the website of the China National Bureau of Statistics. The data includes monthly records spanning from January 2016 to June 2023.

2.1. Consumer Price Index

In addition to the average Consumer Price Index (CPI), we have also collected its components from various departments to enhance the feature set. Table 1 illustrates the structure of the collected data. Each column of Table 1 represents a specific component, such as Food, Tobacco and Alcohol; Clothing; Residential; Daily Necessities and Services; Transportation and Communications; Education, Culture and Entertainment; Healthcare; and Other Goods and Services.

DOI: 10.54254/2754-1169/67/20241290

Due to space limitations, we are unable to display all eight components in Table 1. Each row in Table 1 corresponds to the specific data for the respective month.

Table 1: Data Structure of the CPI (last month = 100).

	Food, Tobacco and Alcohol	Clothing	Residential	Daily Necessities and Services	…	Education, Culture and Entertainment	Healthcare	Other Goods and Services
2023.06	99.4	99.7	100.1	100.6		101.3	100.1	102.4
2023.05	99.8	99.8	100.1	99.5		99.8	100.1	103.1
…								
2016.01	103.6	101.9	101.4	100.6		101.7	102.9	99.6

2.2. Monetary Supply

The Monetary Supply data also includes several detailed components, namely Money and quasi-money (M2) supply at the end of the period (in 100 million yuan), Monetary and quasi-money (M2) supply year-on-year growth (%), Money (M1) supply at the end of the period (in 100 million yuan), Money (M1) supply year-on-year growth (%), Cash in circulation (M0) at the end of the period (in 100 million yuan), and Cash in circulation (M0) year-on-year growth (%). The structure of the Currency Supply data is provided in Table 2, consisting of a total of six features.

When examining the values in Table 2, we observe a significant increase in the currency supply of M2, M1, and M0, reflecting the overall economic growth of China's market.

Table 2: Data Structure of the Monetary Supply.

	M2 Supply	M2 Supply Growth Rate	M1 Supply	M1 Supply Growth Rate	M0 Supply	M0 Supply Growth Rate
2023.06	2873023.83	11.3	695595.48	3.1	105419.20	9.8
2023.05	2820504.68	11.6	675252.98	4.7	104756.71	9.6
…						
2016.01	1416319.55	14.0	412685.64	18.6	72526.51	15.1

3. Algorithm Design

Unlike forecasting the Shanghai Composite Index as time series data [6], our approach involves using monthly index data and the collected CPI, monetary supply data as independent slices. This means that our forecast for the next month's index is calculated based on the current month's index and the data released by the National Bureau of Statistics. We treat this task as a regression problem and utilize Linear Regression as our fundamental algorithm. Furthermore, we enhance our predictions using AdaBoost, which can be considered as a combination of multiple Linear Regression models.

3.1. Linear Regression

Linear regression is a statistical technique utilized to model the relationship between one or more independent variables and a dependent variable. It assumes a linear relationship between these variables, implying that the change in the dependent variable is directly proportional to the change in the independent variable(s). When there is only one independent variable, it is referred to as simple linear regression, while the scenario involving multiple independent variables is known as multiple linear regression.

DOI: 10.54254/2754-1169/67/20241290

Given a dataset $\{y_i, x_{i1}, x_{i2}, \dots, x_{id}\}_{i=1}^{n}$ of n samples, a simple linear regression can be presented as follows:

$$y_i = w_1 x_{i1} + \cdots + w_d x_{id} + b = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{x} + b \tag{1}$$

Where $\boldsymbol{w}$ denotes the weight vector and b is the bias.

3.2. AdaBoost Regression

Given the complexity of real-world data, it becomes challenging to accurately capture the relationship between the dependent variable and the independent variables using simple Linear Regression, particularly in our stock prediction task. Therefore, we have turned to AdaBoost for assistance.

AdaBoost is a learning theory that posits that combining multiple learning algorithms, known as "weak learners", can enhance performance compared to any single algorithm. Let f(x) represent a Linear Regression model. A boosted Linear Regression model takes the form:

$$F_T(x) = \Sigma_{t=1}^{T} \alpha_t f_t(x) \tag{2}$$

Where each $f_t(x)$ is a weak learner that takes a sample x as input and predicts the value of y; α_t refers to the weight of the corresponding learner in the combination, which usually is decided by the error rate e_t of $f_t(x)$ on the dataset as:

$$\alpha_t = \frac{1}{2} \log \frac{1-e_t}{e_t} \tag{3}$$

According to the conclusion in [5], the boosted regressor is assumed to have better performance. Hence, we use it in our prediction task for improvement.

4. Evaluation

In this section, we evaluate the effectiveness of the two factors in predicting the trend of the SSECI. We analyze the performance of the regressor from two perspectives: the precision of predicted values and the accuracy of trend prediction.

4.1. Data Configuration

The feature vectors are composed of the CPI and Monetary Supply values for each month. Additionally, we incorporate the current month's SSECI, specifically the closing price on the last trading day, as auxiliary information. The ground truth for prediction is the SSECI value of the following month.

By adopting this process, we enable independent monthly predictions, mitigating the potential impact of economic cycle fluctuations on the regressor. Moreover, this design facilitates an efficient start for the regressor after training, eliminating the need to collect an extensive dataset over a long period.

4.2. Metrics

We assess the performance of the regressor based on two criteria: the accuracy of predicted values and the precision of trend prediction. The precision is evaluated using the coefficient of determination, denoted as the score, which can be defined as follows:

$$R^2 = 1 - \frac{u}{v} \tag{4}$$

Where $u = \sum_{i \in D} (\tilde{y}_i - y_i)^2$, and $v = \sum_{i \in D} (\bar{y} - y_i)^2$, $\tilde{y}_i$ denotes the predicted value and $\bar{y}$ is the mean of $\{y_i\}_{i \in D}$. The metric of accuracy is simple, which can be formulated as follows:

$$Accuracy = \frac{\sum_{t \in D} I\left((y_{t+1} - y_t) * (\tilde{y}_{t+1} - y_t) > 0\right)}{|D|} \tag{5}$$

Where t denotes the month.

4.3. Performance Analysis

The findings presented in Table 3 illustrate the performance of the LR model and AdaBoost Regressor. It is evident that the AdaBoost Regressor outperforms the LR model in both precision and accuracy. However, both models demonstrate effectiveness in stock prediction.

For instance, both the LR model and AdaBoost Regressor exhibit the ability to accurately capture the trend of the SSECI, with accuracies of 81.48% and 88.89% respectively. This observation indicates a close relationship between the identified factors and China's stock market. Additionally, the precision score of the AdaBoost Regressor (0.7852) surpasses that of the LR model (0.7009), supporting our hypothesis that the combination of multiple LR models can lead to improved performance.

Table 3: Performance of different methods on trend forecast.

	Score	Accuracy
Linear Regression Model	0.7009	81.48%
AdaBoost Regressor	0.7852	88.89%

In conclusion, our study demonstrates that by utilizing sliced features of CPI and monetary supply, we can effectively predict the trend of the SSECI using appropriate machine learning algorithms, such as AdaBoost Regression. This finding suggests that the China's stock market can serve as an indicator of economic conditions, albeit influenced by the government's economic policies.

5. Conclusion

In this paper, we examine two essential sources of data, the Customer Price Index and the Currency Supply, which exhibit a strong correlation with the trending of the China's stock market, specifically the Shanghai Stock Exchange Composite Index. To enhance the predictive capabilities of the Linear Regression algorithm, we propose the utilization of AdaBoost, which has been found to be effective in accurately forecasting the trend of the stock composite index. The experimental results also suggest that the China's stock market can serve as an indicator of economic conditions, albeit influenced by the government's economic policies.

References

[1] *Nazareth, N., & Reddy, Y.V. (2023). Financial applications of machine learning: A literature review. Expert Syst. Appl., 219, 119640.*
[2] *Alzoubi, M. (2022). Stock market performance: Reaction to interest rates and inflation rates. Banks and Bank Systems.*
[3] *Panwar, B., Dhuriya, G., Johri, P., Singh Yadav, S., & Gaur, N.K. (2021). Stock Market Prediction Using Linear Regression and SVM. 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 629-631.*

DOI: 10.54254/2754-1169/67/20241290

[4] *Nathani, N., & Singh, A. (2021). Foundations of Machine Learning. Introduction to AI Techniques for Renewable Energy Systems.*
[5] *Wyner, A. J., Olson, M., Bleich, J., & Mease, D. (2017). Explaining the success of adaboost and random forests as interpolating classifiers. The Journal of Machine Learning Research, 18(1), 1558-1590.*
[6] *Diggs, D.H., & Povinelli, R.J. (2003). A Temporal Pattern Approach for Predicting Weekly Financial Time Series.*